

ISOFORMS OF THE HUMAN VITAMIN D RECEPTOR

Field of the Invention:-

5 The present invention relates to isolated polynucleotide molecules which encode novel isoforms of the human Vitamin D receptor (hVDR) or variant transcripts for hVDR. The polynucleotide molecules may be utilised in, for example, methods of screening compounds for VDR agonists and/or antagonists.

Background of the Invention:-

10 The active hormonal form of vitamin D, 1,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃), has a central role in calcium and phosphate homeostasis, and the maintenance of bone. Apart from these calcitropic effects, 1,25-(OH)₂D₃ has been shown to play a role in controlling cell growth and differentiation in
15 many target tissues. The effects of 1,25-(OH)₂D₃ are mediated by a specific receptor protein, the vitamin D receptor (VDR), a member of the nuclear receptor superfamily of transcriptional regulators which also includes steroid, thyroid and retinoid receptors as well as a growing number of orphan receptors. Upon binding hormone the VDR regulates gene expression by
20 direct interaction with specific sequence elements in the promotor regions of hormone responsive target genes. This transactivation or repression involves multiple interactions with other protein cofactors, heterodimerisation partners and the transcription machinery.

25 Although a cDNA encoding the human VDR was cloned in 1988 (1), little has been documented characterising the gene structure and pattern of transcription since that time. The regulation of VDR abundance is one potentially important mechanism for modulating 1,25-(OH)₂D₃ responsiveness in target cells. It is also possible that VDR has a role in non-transcriptional pathways, perhaps via localization to a non-nuclear
30 compartment and/or interaction with components of other signalling pathways. However, the question of how VDRs are targetted to different cell types and how they are regulated remains unresolved. There have been many reports in the literature describing translational or transcriptional control of VDR levels, both homologously and heterologously, mostly in non-human
35 systems.

A recent study (2) showed that in the kidney, alternative splicing of human VDR transcripts transcribed from a GC rich promotor generates several transcripts which vary only in their 5' UTRs. The present inventors have now identified further upstream exons of the VDR gene which generate 5' variant transcripts, suggesting that the expression of the VDR gene is regulated by more than one promoter. A subset of these transcripts is expressed in a restricted tissue-specific pattern and further variant transcripts have the potential to encode an N-terminally variant protein. These results may have implications for understanding the actions of 1,25-(OH)₂D₃ in different tissues and cell types, and the possibility that N-terminally variant VDR proteins may be produced has implications for altered activities such as transactivation function or subcellular localisation of the receptor protein. Furthermore, these variants, by their level, tissue specificity, subcellular localisation and functional activity, may yield targets for pharmaceutical intervention. The variants may also be useful in screening potential analogs and/or antagonists of vitamin D compounds.

Disclosure of the Invention:-

In a first aspect, the invention provides an isolated polynucleotide molecule encoding a human Vitamin D receptor (hVDR) isoform, said polynucleotide molecule comprising a nucleotide sequence which includes sequence that substantially corresponds or is functionally equivalent to that of exon 1d of the human VDR gene.

Exon 1d (referred to as exon 1b in the Australian Provisional Patent Specification No. PO9500) is a 96 bp exon located 296 bp downstream from exon 1a (2). The sequence of exon 1d is:

5'GTTTCCTTCTTCTGTCGGGGCGCCTTGGCATGGAGTGGAGGAATAAGAA
AAGGAGCGATTGGCTGTCGATGGTGCTCAGAACTGCTGGAGTGGAGG3'
(SEQ ID NO: 1).

The nucleotide sequence of the polynucleotide molecule of the first aspect of the invention, preferably does not include sequence corresponding to that of exon 1a, exon 1f and/or exon 1e. However, the nucleotide sequence of the polynucleotide molecule of the first aspect of the invention, may or

may not include sequence that substantially corresponds or is functionally equivalent to that of exon 1b and/or exon 1c.

Preferably, the polynucleotide molecule of the first aspect comprises a nucleotide sequence which includes;

5 (i) sequence that substantially corresponds or is functionally equivalent to that of exons 1d, 1c and 2-9 and encodes a VDR isoform of approximately 477 amino acids,

(ii) sequence that substantially corresponds or is functionally equivalent to that of exons 1d and 2-9 and encodes a VDR isoform of
10 approximately 450 amino acids, or

(iii) sequence that substantially corresponds or is functionally equivalent to that of exons 1d and 2-9 and further includes a 152 bp intronic sequence, and encodes a truncated VDR isoform of approximately 72 amino acids.

15 Most preferably, the polynucleotide molecule of the first aspect of the invention comprises a nucleotide sequence substantially corresponding to that shown as SEQ ID NO: 2, SEQ ID NO: 3 or SEQ ID NO: 4.

In a second aspect, the invention provides an isolated polynucleotide molecule encoding a human Vitamin D receptor (hVDR), said polynucleotide
20 molecule comprising a nucleotide sequence which includes sequence that substantially corresponds to that of exon 1f and/or 1e of the human VDR gene.

Exon 1f is a 207bp exon located more than 9kb upstream from exon 1a (2) bp upstream from exon 1c(8). The sequence of exon 1f is:

25

5'TGCGACCTTGGCGGTGAGCCTGGGGACAGGGGTGAGGC
CAGAGACGGACGGACGCAGGGGCCCCGCCCCAAGGCGAGGG
AGAACAGCGGCACTAAGGCAGAAAGGAAGAGGGCGGTGTG
TTCACCCGCAGCCCAATCCATCACTCAGCAACTCCTAGAC
30 GCTGGTAGAAAGTTCCTCCGAGGAGCCTGCCATCCAGTCGT
CCGTGCAG3' (SEQ ID NO: 5)

Exon 1e is a 157 bp exon located 1826bp upstream from exon 1a (2).
The sequence of exon 1e is:

5'AGGCAGCATGAAACAGTGGGATGTGCAGAG
AGAAGATCTGGGTCCAGTAGCTCTGACACTCCTCAGCTGT
AGAAACCTTGACAACTCTGCACATCAGTTGTACAATGGAA
CGGTATTTTTTACTCTTCATGTCTGAAAAGGCTATGATAA
AGATCAA3' (SEQ ID NO: 6)

5
10 The nucleotide sequence of the polynucleotide molecule of the second aspect of the invention, preferably does not include sequence corresponding to that of exon 1a, 1d or 1b. However, the nucleotide sequence of the polynucleotide molecule of the second aspect of the invention, may or may not include sequence that substantially corresponds or is functionally equivalent to that of exon 1c.

15 Preferably, the nucleotide molecule of the second aspect comprises a nucleotide sequence which includes sequence that substantially corresponds or is functionally equivalent to that of exons 1f and 2-9.

Most preferably, the polynucleotide molecule of the first aspect of the invention comprises a nucleotide sequence substantially corresponding to that shown as SEQ ID NO: 7.

20 The polynucleotide molecule of the first or second aspects may be incorporated into plasmids or expression vectors (including viral vectors), which may then be introduced into suitable host cells (e.g. bacterial, yeast, insect and mammalian host cells). Such host cells may be used to express the VDR or functionally equivalent fragment thereof encoded by the isolated polynucleotide molecule.

25 Accordingly, in a third aspect, the present invention provides a host cell transformed with the polynucleotide molecule of the first or second aspect.

30 In a fourth aspect, the present invention provides a method of producing a VDR or a functionally equivalent fragment thereof, comprising culturing the host cell of the first or second aspect under conditions enabling the expression of the polynucleotide molecule and, optionally, recovering the VDR or functionally equivalent fragment thereof.

35 Preferably, the host cell is of mammalian origin. Preferred examples include NIH 3T3 and COS 7 cells.

In a preferred embodiment, the VDR or functionally equivalent fragment thereof is localised to a cell membrane or other subcellular compartment as distinct from a nuclear localisation.

5 The polynucleotide molecules of the first aspect of the invention encode novel VDR isoforms which may be of interest both clinically and commercially. By using the polynucleotide molecule of the present invention it is possible to obtain VDR isoform proteins or functionally equivalent fragments thereof in a substantially pure form.

10 Accordingly, in a fifth aspect, the present invention provides a human VDR isoform or functionally equivalent fragment thereof encoded by a polynucleotide molecule of the first aspect, said VDR isoform or functionally equivalent fragment thereof being in a substantially pure form.

15 In a sixth aspect, the present invention provides an antibody or antibody fragment capable of specifically binding to the VDR isoform of the fourth aspect.

The antibody may be monoclonal or polyclonal, however, it is presently preferred that the antibody is a monoclonal antibody. Suitable antibody fragments include Fab, F(ab')₂ and scFv.

20 In an eighth aspect, the present invention provides a non-human animal transformed with a polynucleotide molecule according to the first or second aspect of the invention.

25 In a seventh aspect, the invention provides a method for detecting agonist and/or antagonist compounds of a VDR isoform of the fourth aspect, comprising contacting said VDR isoform, functionally equivalent fragment thereof or a cell transformed with and expressing the polynucleotide molecule of the first aspect, with a test compound under conditions enabling the activation of the VDR isoform or functionally equivalent fragment thereof, and detecting an increase or decrease in the activity of the VDR isoform or functionally equivalent fragment thereof.

30 An increase or decrease in activity of the receptor or functionally equivalent fragment thereof may be detected by measuring changes in interactions with known cofactors (e.g. SRC-1, GRIP-1 and TFIIB) or unknown cofactors (e.g. through use of the yeast dual hybrid system).

35 In a ninth aspect, the present invention provides an oligonucleotide or polynucleotide probe comprising a nucleotide sequence of 10 or more nucleotides, the probe comprising a nucleotide sequence such that the probe

specifically hybridises to the polynucleotide molecule of the first or second aspect under high stringency conditions (Sambrook *et al.*, *Molecular Cloning: a laboratory manual*, Second Edition, Cold Spring Harbor Laboratory Press).

Preferably, the probe is labelled.

5 In a tenth aspect, the present invention provides an antisense polynucleotide molecule comprising a nucleotide sequence capable of specifically hybridising to an mRNA molecule which encodes a VDR encoded by the polynucleotide molecule of the first or second aspect, so as to prevent translation of the mRNA molecule.

10 Such antisense polynucleotide molecules may include a ribozyme region to catalytically inactivate mRNA to which it is hybridised.

The polynucleotide molecule of the first or second aspect of the invention may be a dominant negative mutant which encodes a gene product causing an altered phenotype by, for example, reducing or eliminating the activity of endogenous VDR.

15 In an eleventh aspect, the invention provides an isolated polynucleotide molecule comprising a nucleotide sequence substantially corresponding or, at least, showing >75% (preferably >85% or, even more preferably, >95%) sequence identity to:

20 (i) 5'TGCGACCTTGCGGCTGAGCCTGGGGACAGGGGTGAGGCCAGAGA
CGGACGGACGCAGGGGCCCCGCCCCAAGGCGAGGGAGAACAGCGGCACTA
AGGCAGAAAGGAAGAGGGCGGTGTGTTACCCCGCAGCCCAATCCATCAC
TCAGCAACTCCTAGACGCTGGTAGAAAGTTCCTCCGAGGAGCCTGCCATC
25 CAGTCGTGCGTGCGAG 3'(exon 1f) (SEQ ID NO: 5),

(ii) 5'AGGCAGCATGAAACAGTGGGATGTGCAGAGAGAAGATCTGGGTC
CAGTAGCTCTGACACTCCTCAGCTGTAGAAACCTTGACAACCTCTGCACAT
CAGTTGTACAATGGAACGGTATTTTTTACTCTTCATGTCTGAAAAGGCTA
30 TGATAAAGATCAA3' (exon 1e) (SEQ ID NO: 6), or

(iii) 5'GTTTCCTTCTTCTGTCTGGGGCGCCTTGGCATGGAGTGGAGGAATA
AGAAAAGGAGCGATTGGCTGTCTGATGGTGCTCAGAACTGCTGGAGTGGAG
GG3' (exon 1d) (SEQ ID NO: 1).

The polynucleotide molecules of the eleventh aspect may be useful as probes for the detection of VDR variant transcripts and as such may be useful in assessing cell or tissue-specific expression of variant transcripts.

The terms "substantially corresponds" and "substantially corresponding" as used herein in relation to nucleotide sequences is intended to encompass minor variations in the nucleotide sequence which due to degeneracy in the DNA code do not result in a substantial change in the encoded protein. Further, this term is intended to encompass other minor variations in the sequence which may be required to enhance expression in a particular system but in which the variations do not result in a decrease in biological activity of the encoded protein.

The term "functionally equivalent" as used herein in relation to nucleotide sequences encoding a VDR isoform is intended to encompass nucleotide sequence variants of up to 5% sequence divergence (i.e. retaining 95% or more sequence identity) which encode VDR isoforms of substantially equivalent biological activity(ies) as said VDR isoform.

The term "functionally equivalent fragment" as used herein in respect of a VDR isoform is intended to encompass functional peptide and polypeptide fragments of said VDR isoform which include the domain or domains which bestow the biological activity characteristic of said VDR isoform.

The terms "comprise", "comprises" and "comprising" as used throughout the specification are intended to refer to the inclusion of a stated step, component or feature or group of steps, components or features with or without the inclusion of a further step, component or feature or group of steps, components or features.

The invention will hereinafter be further described by way of the following non-limiting example and accompanying figures.

Brief description of the figures:-

FIG.1. (A) Human VDR gene locus. Four overlapping cosmid clones were isolated from a human lymphocyte genomic library (Stratagene) and directly sequenced. Clone J5 extends from the 5' flanking region to intron 2; AE, from intron 1b to intron 5; D2, from intron 3 to the 3' UTR; WE, from intron 6 through the 3' flanking region. Sequence upstream of exon 1f was obtained by

anchored PCR from genomic DNA. (B) Structure of hVDR transcripts.

Transcripts 1–5 originate from exon 1a. Transcript 1 corresponds to the published cDNA (1). Transcripts 6–10 originate from exon 1d and transcripts 11–14 originate from exon 1f. Boxed numbers indicate the major transcript (based on the relative intensities of the multiple PCR products) within each exon-specific group of transcripts generated with a single primer set. While all transcripts have a translation initiation codon in exon 2, exon 1d transcripts have the potential to initiate translation upstream in exon 1d, with transcripts 6 and 9 encoding VDR proteins with extended N termini. (C) N-terminal variant proteins encoded by novel hVDR transcripts. Transcript 1 corresponds to the published cDNA sequence (1) and encodes the 427-aa hVDR protein. Transcripts 6 and 9 code for a protein with an extra 50 aa or 23 aa, respectively, at the N-terminal. The 23 aa of the hVDR A/B domain are shown in bold.

FIG. 2. RT-PCR analysis of expression of variant hVDR transcripts. (A) Exon 1a transcripts (220 bp, 301 bp, 342 bp, 372 bp, and 423 bp). (B) Exon 1d transcripts (224 bp, 305 bp, 346 bp, 376 bp, and 427 bp). (C) Exon 1f transcripts (228 bp, 309 bp, 387 bp, and 468 bp). RT-PCR was carried out with exon 1a-, 1d-, or 1f-specific forward primers and a common reverse primer in exon 3. The sizes of the PCR products and the pattern of bands are similar in A and B by virtue of the identical splicing pattern of exon 1a and 1d transcripts and the fact that primers were designed to generate PCR products of comparable sizes. All tissues and cell lines are human in origin.

FIG. 3. Functional analysis of sequence-flanking exons 1a and 1d (A) and exon 1f (B) in NIH 3T3 (solid bars) and COS 7 cells (open bars).

The parent vector pGL3basic was used as a promoterless control, and a promoter-chloramphenicol acetyltransferase (CAT) gene reporter construct was cotransfected as an internal control for transfection efficiency in each case. The activity of each construct was corrected for transfection efficiency and for the activity of the pGL3basic empty vector control and expressed as a percentage of the activity of the construct 1a(-488,+75) SEM of at least three separate transfections. Exon 1a and 1d flanking constructs are defined in relation to the transcription start site of exon

1a, designated 11, which lies 54 nt upstream of the published cDNA (1). Exon 1f flanking constructs are defined relative to the exon 1f transcription start site, designated 11. Transcription start sites were determined by the 5' termini of the longest RACE clones. The open box corresponds to the GC-rich region.

FIG 4. Provides the nucleotide sequence of novel exons detected by 5' RACE: (A) exon 1b (SEQ ID NO: 8), (B) exon 1f (SEQ ID NO: 5) [P1f is indicated by an arrow above the sequence], (C) exon 1e (SEQ ID NO: 6), (D) exon 1d (SEQ ID NO: 1) [in-frame ATG codons are highlighted and P1d is indicated by an arrow above the sequence]. Intronic sequences are shown in lower case. Canonical splice site consensus sequences are indicated in bold. The transcription start sites for exons 1f and 1d were determined by the 5' termini of RACE clones. No intron sequence is shown 3' to exon 1f as cosmid clone J5 terminated in the intron between exons 1f and 1e.

FIG 5. Provides the nucleotide sequence corresponding to transcript 6 (see figure 1) (SEQ ID NO: 2), together with the predicted amino acid sequence (SEQ ID NO: 9) of the encoded protein. Nucleotides 1-96 correspond to exon 1d; nucleotides 97-1463 correspond to exons 1c to the stop codon in exon 9 (or nucleotides -83-1283 of the hVDR cDNA (1)).

FIG 6. Provides the nucleotide sequence corresponding to transcript 9 (see figure 1) (SEQ ID NO: 3), together with the predicted amino acid sequence (SEQ ID NO: 10) of the encoded protein. Nucleotides 1-96 correspond to exon 1d; nucleotides 97 - 1382 correspond to exon 2 to the stop codon in exon 9 (or nucleotides -2 - 1283 of the hVDR cDNA (1)).

FIG 7. Provides the nucleotide sequence corresponding to transcript 10 (see figure 1) (SEQ ID NO: 4), together with the predicted amino acid sequence (SEQ ID NO: 11) of the encoded protein. Nucleotides 1-96 correspond to exon 1d; nucleotides 97-244 correspond to exon 2; nucleotides 245-396 correspond to intronic sequence immediately 3' to exon 2; nucleotides 397-1534 correspond to exons 3 to the stop codon in exon 9 (or nucleotides 146-1283 of the hVDR cDNA (1)).

FIG 8. Provides the nucleotide sequence corresponding to transcript 11 (see figure 1) (SEQ ID NO: 7), together with the predicted amino acid sequence (SEQ ID NO: 12) of the encoded protein. Nucleotides 1-207 correspond to exon 1f; nucleotides 208-1574 correspond to exon 1c to the stop codon in exon 9 (or nucleotides -83-1283 of the hVDR cDNA (1)).

Example:-

EXPERIMENTAL PROCEDURES

Isolation and Characterisation of Genomic Clones

A human lymphocyte cosmic library (Stratagene, La Jolla, Ca) was screened using a 2.1kb fragment of the hVDR cDNA encompassing the entire coding region but lacking the 3'UTR, a 241 bp PCR product spanning exons 1 to 3 of the human VDR cDNA, and a 303 bp PCR product spanning exons 3 and 4 of the hVDR cDNA, following standard colony hybridisation techniques. DNA probes were labelled by nick translation (Life Technologies, Gaithersburg, MD) with [α^{32} P] dCTP. Positively hybridising colonies were picked and secondary and tertiary screens carried out until complete purification. Cosmid DNA from positive clones was purified (Qiagen), digested with different restriction enzymes and characterised by Southern blot analysis using specific [γ^{32} P]ATP labelled oligonucleotides as probes. Cosmid clones were directly sequenced using dye-termination chemistry and automated fluorescent sequencing on an ABI Prism, 377 DNA Sequencer (Perkin-Elmer, Foster City, Ca). Sequence upstream of the most 5' cosmid was obtained by anchored PCR from genomic DNA using commercially available anchor ligated DNA (Clontech, Palo Alto, Ca).

Rapid Amplification of cDNA 5-prime Ends (5'-RACE)

Alternative 5' variants of the human VDR gene were identified by 5'RACE using commercially prepared anchor-ligated cDNA (Clontech) following the instructions of the manufacturer. Two rounds of PCR using nested reverse primers in exons 3 and 2 (P 1: 5'ccgcttcattgcttcgcctgaagaagcc-3', P2: 5'-tgcagaattcacaggtcatagcattgaag-3') were carried out on a Corbett FTS-4000 Capillary Thermal Sequencer (Corbett Research, NSW, Australia). After 26 cycles of PCR, 2% of the primary reaction was reamplified for 31 cycles.

The PCR products were cloned into PUC18 and sequenced by the dideoxy chain termination method.

Cell-Culture

5 The embryonal kidney cell line, HEK-293, an embryonic intestine cell line, Intestine-407 and WS 1, a foetal skin fibroblast cell line were all cultured in Eagle's MEM with Earle's BSS and supplemented with either 10% heat-inactivated FBS, 15% FBS or 10% FBS with non-essential amino acids, respectively. The osteosarcoma cell lines MG-63 and Saos-2 were cultured in
10 Eagle's MEM with nonessential amino acids and 10% heat-inactivated FBS and McCoy's 5a medium with 15% FBS, respectively. The breast carcinoma cell line T47D and the colon carcinoma cell lines LIM 1863 and COLO 206F were cultured in RPMI medium supplemented with 0.2 IU bovine insulin/ml and 10% FBS, 5% FBS or 10% FBS, respectively. LIM 1863 were a gift from
15 R.H. Whitehead (3). HK-2 kidney proximal tubule cells were grown in keratinocyte-serum free medium supplemented with 5ng/ml recombinant EGF, 40ug/ml bovine pituitary extract. BC1 foetal osteoblast-like cells were kindly donated by R. Mason (4) and were grown in Eagle's MEM with 5% FBS and 5mg/L vitamin C. Unless otherwise stated all cell lines were obtained
20 from the American Type Culture Collection (Manassas, VA).

Reverse Transcriptase-PCR (RT-PCR).

Total RNA extracted from approximately 1.5×10^3 cells. from leukocytes prepared from 40 ml blood, or from human tissue using acid-
25 phenol extraction was purified by using a guanidium isothiocyanate-caesium chloride step gradient. First-strand cDNA was synthesized from 5 µg of total RNA primed with random hexamers (Promega) using Superscript II reverse transcriptase (Life Technologies). One-tenth of the cDNA (2µl) was used for subsequent PCR, with 36 cycles of amplification, using exon-specific forward
30 primers (exon 1a: corresponding to nucleotides 1-21 of hVDR cDNA (1); exon 1d: 5'-GGCTGTCGATGGTGCTCAGAAC-3'; exon 1f: 5'-AAGTTCCTCCGAGGAGCCTGCC-3'); and a common reverse primer in exon 3 [corresponding to nucleotides 301-280 of hVDR cDNA (1)]. All RT-PCRs were repeated multiple times by using
35 RNA/cDNA prepared at different times from multiple sources. Each PCR included an appropriate cDNA-negative control, and additional controls

included RT-negative controls prepared alongside cDNA and RNA/cDNA prepared from VDR-negative cell lines. PCR products were separated on 2% agarose and visualized with ethidium bromide staining.

5 *Functional Analysis of hVDR Gene Promoters.*

- Sequences flanking exons 1a, 1d, and 1f (see Fig. 1A) were PCR-amplified by using Pfu polymerase (Stratagene) and cloned into the pGL3basic vector (Promega) upstream of the luciferase gene reporter. Promoter-reporter constructs were transfected into NIH 3T3 and COS 7 cells by using the standard calcium phosphate-precipitation method. Cells were seeded at $2.3 \pm 2.5 \times 10^6$ per 150-cm² flask the day before transfection. Several hours before the precipitates were added the medium was changed to DMEM with 2% charcoal-stripped FBS. Cells were exposed to precipitate for 16 h before subculturing and were harvested 24 h later. The parent vector pGL3basic was used as a promoterless control in these experiments and a simian virus 40 promoter-chloramphenicol acetyltransferase (CAT) gene reporter construct was cotransfected as an internal control for transfection efficiency in each case. The activity of each construct was corrected for transfection efficiency and for the activity of the pGL3 basic empty vector control and expressed as a percentage of the activity of the construct 1a(-488,+75). Luciferase and CAT assays were carried out in triplicate, and each construct was tested in transfection at least three times.

25 *RESULTS*

25 *Identification of Alternative 5' Variants of the hVDR Gene.*

- Upstream exons were identified in human kidney VDR transcripts by 5' RACE (exons 1f, 1e, 1d, and 1b) and localized by sequencing of cosmid clones (Fig. 1A). To verify these results and to characterize the structure of the 5' end of the VDR gene, exon-specific forward primers were used with a common reverse primer in exon 3 to amplify specific VDR transcripts from human tissue and cell line RNA (Fig. 1B). The identity of these PCR products was verified by Southern blot and by cloning and sequencing. Five different VDR transcripts originating from exon 1a were identified. The major transcript (transcript 1 in Fig. 1B) corresponds to the published cDNA sequence (1). Three less-abundant forms (2, 3, and 4 in Fig. 1B) arise from

alternative splicing of exon 1c and a novel 122-bp exon 1b into or out of the final transcript. These three variant transcripts were described recently by Pike and colleagues (2). A fifth minor variant was identified (5 in Fig. 1B) that lacks exons 1b and 1c, but includes an extra 152 bp of intronic sequence immediately 3' to exon 2, potentially encoding a truncated protein as a result of an in-frame termination codon in intron 2.

Four more transcripts were characterized that originate from exon 1f, a novel 207-bp exon more than 9 kb upstream from exon 1a. The major 1f-containing transcript (11 in Fig. 1B) consists of exon 1f spliced immediately adjacent to exon 1c. Three less-abundant variants (12, 13, and 14 in Fig. 1B) arise from alternative splicing of exon 1c and a novel 159-bp exon 1e into or out of the final transcript. All these hVDR variants differ only in their 5' UTRs and encode identical proteins from translation initiation in exon 2.

Of considerable interest, another five hVDR transcripts were identified that originate from exon 1d, a novel 96-bp exon located 296 bp downstream from exon 1a. The major exon 1d-containing transcript (6 in Fig. 1B) utilizes exon 1d in place of exon 1a of the hVDR cDNA. Three minor variants (7, 8, and 9 in Fig. 1B) arise from alternative splicing of exons 1b and 1c into or out of the transcript, analogous to the exon 1a-containing variants 2, 3, and 4. A fifth minor variant transcript (10 in Fig. 1B) lacks exons 1b and 1c, but includes 152 bp of intron 2 analogous to the exon 1a-containing transcript 5, and also potentially encodes a truncated protein. Two of these exon 1d-containing hVDR transcripts encode an N-terminal variant form of the hVDR protein. Utilization of an ATG codon in exon 1d, which is in a favorable context and in-frame with the major translation start site in exon 2, would generate a protein with an additional 50 aa N-terminal to the ATG codon in exon 2 in the case of variant 6 or 23 aa in the case of variant 9 (Fig. 1C).

The relative level of expression of the different transcripts is difficult to address with PCR since relatively minor transcripts may be amplified. However, Southern blots of PCR products from the linear range of PCR amplification indicated that equivalent amounts of PCR product were accumulated after 26 cycles for exon 1a transcripts compared with 30 cycles for exon 1d transcripts, suggesting that 1d abundance is about 5% of that of 1a transcripts. This is consistent with the frequency of clones selected and sequenced from RACE analysis of two separate samples of kidney RNA: 1a (21/27; 78%), 1d (2/27; 7%), and 1f (4/27; 15%). RT-PCR with exon 1a-, 1d-, or

1f-specific forward primers and reverse primers in exons 7 or 9, followed by cloning and sequencing, suggests that these 5' variant transcripts are not associated with differences at the 3' end of the transcript.

5 *Exon-Intron Organization of the hVDR Gene.*

Overlapping cosmid clones were isolated from a human lymphocyte genomic library and characterized by hybridization to exon-specific oligonucleotide probes (Fig. 1A). The exon-intron boundaries of the hVDR gene were determined by comparison of the genomic sequence from cosmid clones with the cDNA sequence. Upstream exons were localized in the VDR gene by sequencing cosmid clones, which extend approximately 7 kb into the intron between exons 1e and 1f, enabling verification of both their sequence and the presence of consensus splice donor/acceptor sites. Sequence upstream of exon 1f was obtained by anchored PCR from genomic DNA by using commercially available anchor-ligated DNA (CLONTECH). In total, the hVDR gene spans more than 60 kb and consists of at least 14 exons (Fig. 1A).

15 *Tissue-Specific Expression of hVDR Transcripts.*

The pattern of expression of variant hVDR transcripts was examined by RT-PCR in a variety of cell lines and tissues with exon 1a-, 1d-, or 1f-specific forward primers and a common reverse primer in exon 3. Exon 1a and 1d transcripts (Fig. 1B, variants 1-10) were coordinately expressed in all RNA samples analyzed (Fig. 2 A and B). Exon 1f transcripts (Fig. 1B, variants 11-14), however, were detected only in RNA from human kidney tissue (two separate samples), human parathyroid adenoma tissue, and an intestinal carcinoma cell line, LIM 1863 (Fig. 2C). Interestingly, these represent major target tissues for the calcitropic effects of vitamin D.

20 *Functional Analysis of hVDR Gene Promoters.*

Promoter activities of the 5' flanking regions of exons 1a, 1d, and 1f were examined in NIH 3T3 and COS 7 cells (Fig. 3). Sequences flanking exon 1a exhibited high promoter activity in both cell lines (Fig. 3A). Maximum luciferase expression of 36- and 54-fold over the empty vector was attained for construct 1a(-488, +75) in NIH 3T3 and COS 7 cells, respectively. This activity could be attributed largely to a GC-rich region containing multiple consensus Sp1-binding motifs lying within 100 bp immediately adjacent to

the transcription start site. This region alone, upstream of a luciferase reporter [construct 1a(-94,+75)], accounted for 43% of the maximum activity observed in NIH 3T3 cells and 86% of the maximum observed in COS 7 cells. The removal of this GC-rich region [construct 1a(-29,+75)] reduced luciferase activity to only 13% of the maximum in NIH 3T3 and 19% in COS 7 cells. Despite the fact that VDR transcripts that originated from exon 1d were identified, distinct promoter activity was not associated with sequences within 300 bp of exon 1d [constructs 1d(+87,+424) and 1d(+244,+424)]; rather, the sequence immediately adjacent to exon 1d may contain a suppressor element (Fig. 3A). Construct 1a-1d(-846,+470), spanning the 5' flanking regions of both exons 1a and 1d, resulted in only 42% and 60% of the activity of 1a(-898,+75) in NIH 3T3 and COS 7 cells, whereas the 3' deletion of 227 bp restored luciferase activity to 65% and 97% of the activity of 1a(-898,+75), respectively. Similarly, the 5' truncated construct 1a-1d (-94,+470), spanning the 5' flanking regions of both 1a and 1d, resulted in only 35% and 40% of the activity of 1a(-94,+75), while a further 3' deletion of 227 bp restored luciferase activity to 69% and 91% of the activity of 1a(-94,+75) in NIH 3T3 and COS 7 cells. It is possible that transcription from exons 1a and 1d is driven by overlapping promoter regions rather than from two distinct promoters, as has been described for the mouse androgen receptor gene.

Sequence upstream of exon 1f showed significant promoter activity in NIH 3T3 cells of 22% of that of the most active construct, 1a(-488,+75), or 9-fold over pGL3basic [construct 1f(-1168,+58)] (Fig. 3B). A shorter construct [1f(-172,+58)] had similar activity, with evidence of a suppressor element (between nucleotides -278 and +172) able to repress luciferase activity by 70%. Interestingly, the same constructs were not active in COS 7 cells. This cell line-specific activity of exon 1f flanking sequences may reflect a requirement for tissue- or cell-specific protein factors.

Identification of VDR isoforms in whole cell lysates

The existence of a VDR isoform including exons 1d and 1c has been confirmed in cell lysates from multiple human, monkey, rat and mouse cell lines derived from kidney, intestine, liver and bone, by immunoprecipitation (using the anti-VDR 9A7 rat monoclonal antibody; Affinity Bioreagents Inc.,

Golden, Colorado) followed by Western blot analysis. The 1d- and 1c-exon-specific antibodies detected the same band in all immunoprecipitations.

DISCUSSION

5 The present inventors have identified 5' variant transcripts of the hVDR that suggest the existence of alternative promoters. These transcripts may not have been discriminated in previous Northern analyses because of their similarity in size. Transcription initiation from exons 1a or 1f and
10 alternative splicing generate VDR transcripts that vary in their 5' UTRs but encode the same 427-aa protein. Transcription initiation from exon 1d and alternative splicing generate hVDR transcripts with the potential to encode variant proteins with an additional 50 or 23 aa at the N terminus. There was
15 no evidence that these 5' variants are associated with differences at the 3' end of the transcript. Although isoforms are common in other members of the nuclear receptor superfamily, the only evidence for isoforms of the hVDR is a common polymorphism in the triplet encoding the initiating methionine of the 427-aa form of the VDR that results in initiation of translation at an
20 alternative start codon beginning at the 10th nucleotide down-stream, encoding a protein truncated by 3 aa at the N terminus (5). Similarly, two forms of the avian VDR, differing in size by 14 aa, are generated from a single transcript by alternative translation initiation (6), and in the rat a dominant-negative VDR is generated by intron retention (7).

Heterogeneity in the 5' region is a common feature of other nuclear
25 receptor genes. Tissue-specific alternative-promoter usage generates multiple transcripts of the human estrogen receptor α (ER α), the human and rat mineralocorticoid receptors, and the mouse glucocorticoid receptor (GR), which differ in their 5' UTRs but code for identical proteins. However, other
30 members of the nuclear receptor superfamily have multiple, functionally distinct isoforms arising from differential promoter usage and/or alternative splicing. The generation of N-terminal variant protein isoforms has been described for the progesterone receptor (PR), peroxisome proliferator-activated receptor (PPAR α), and the retinoid and thyroid receptors. Some
35 receptor isoforms exhibit differential promoter-specific transactivation activity. The N-terminal A/B regions of many nuclear receptor proteins possess a ligand-independent transactivation function (AF1). An AF1

domain has been demonstrated for the thyroid receptor b1 (TRb1), ER, GR, PR, PPAR γ , and the retinoid receptors. The activity of the AF1 domain has been shown to vary in both a tissue- and promoter-specific manner. The N-terminal A/B region of nuclear receptors is the least-conserved domain across the family and between receptor subtypes, varying considerably both in length and sequence. The VDR, however, is unusual as its N-terminal A/B region is much shorter than that of other nuclear receptors, with only 23 aa N-terminal to the DNA-binding domain, and deletion of these residues seems to have no effect on VDR function. This region in other receptors is associated with optimal ligand-dependent transactivation and can interact directly with components of the basal transcription complex. Two stretches of basic amino acid residues, RNKKR and RPHRR, in the predicted amino acid sequences of the variant hVDR N termini (Fig. 1C) resemble nuclear localization signals. An N-terminal variant VDR protein therefore might exhibit different transactivation potential, possibly mediated by different protein interactions, or may specify a different subcellular localization. The tissue-specific expression of exon 1f-containing transcripts is mediated by a distal promoter more than 9 kb upstream of exons 1a and 1d. Exon 1f transcripts were detected only in kidney tissue, parathyroid adenoma tissue, and an intestinal cell line, LIM 1863. It is interesting that these tissues represent major target tissues for the calcitropic effects of vitamin D. The absence of 1f-containing transcripts in two other kidney cell lines, HK-2 (proximal tubule) and HEK-293 (embryonal kidney), as well as one other embryonal intestinal cell line, Intestine-407, suggests that the expression of 1f transcripts is cell type-specific. The cell line-specific activity of exon 1f flanking sequences in promoter reporter assays may reflect a requirement for tissue- or cell-specific protein factors to mediate expression from this promoter.

This study has demonstrated that expression of the human VDR gene, which spans more than 60 kb and consists of 14 exons, is under complex transcriptional control by multiple promoters. The expression of multiple exon 1f transcripts is mediated by utilization of a distal tissue-specific promoter. Transcription from a proximal promoter, or promoters, generates multiple variant hVDR transcripts, two of which code for N-terminal variant proteins. Multiple, functionally distinct isoforms mediate the tissue- and/or developmental-specific effects of many members of the nuclear receptor

5 superfamily. Although the actual relative abundance of the various transcripts and their levels of translation *in vivo* have not yet been characterized, the results suggest that major variant isoforms of the hVDR exist. Differential regulation of these hVDR gene promoters and of alternative splicing of variant VDR transcripts may have implications for understanding the various actions of $1,25-(OH)_2D_3$ in different cell types, and variant VDR transcripts may play a role in tissue specific VDR actions in bone and calcium homeostasis.

10 It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to
15 be considered in all respects as illustrative and not restrictive.

005100 28105100

References:-

1. Baker, A.R. et.al. (1988) Proc. Natl. Acad. Sci. USA 85. 3294-3298
2. Miyamoto, K. et.al. (1997) Mol. Endocrinol. 11, 1165-1179
- 5 3. Whitehead, R.H.et.al. (1987) Cancer Res. 47, 2683-2689
4. Slater, M. et al. (1994) Am. J. Physiology 267, E990-1001
5. Saijo, T.et.al. (1991) Amj.Hum.Genet. 49, 668-673
6. Lu, Z. et.al. (1997) Arch. Biochem. Biophys. 339. 99-106
7. Ebihara, K. et.al. (1996) Mol.Cell.Biol. 16, 3393-3400

005150 28160560